Leveraging large-scale physics-based simulations to improve immediate response after earthquakes

Xavier Bellagamba[1] and Brendon A. Bradley[1]
[1]University of Canterbury
xavier.bellagamba@pg.canterbury.ac.nz

QuakeCoRE
NZ Centre for Earthquake Resilience
Te Hiranga Rū

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Motivation

The immediate response from emergency managers, authorities, and insurances following a major earthquake is based on the best estimate of the regional ground motion (GM) intensity. Current methods assume earthquakes originating from a point-source (Fig. 1a). Compared to the closest approximation that took months to develop (Fig. 1b), the point-source estimation yields relatively poor results. As earthquakes are rare events, so is their associated ground truth data. However, hazard analyses are carried out with physics-based simulations that can be used as a substitute. A machine-learning-based solution is proposed to enhance the first GM estimate that leverages these simulations.

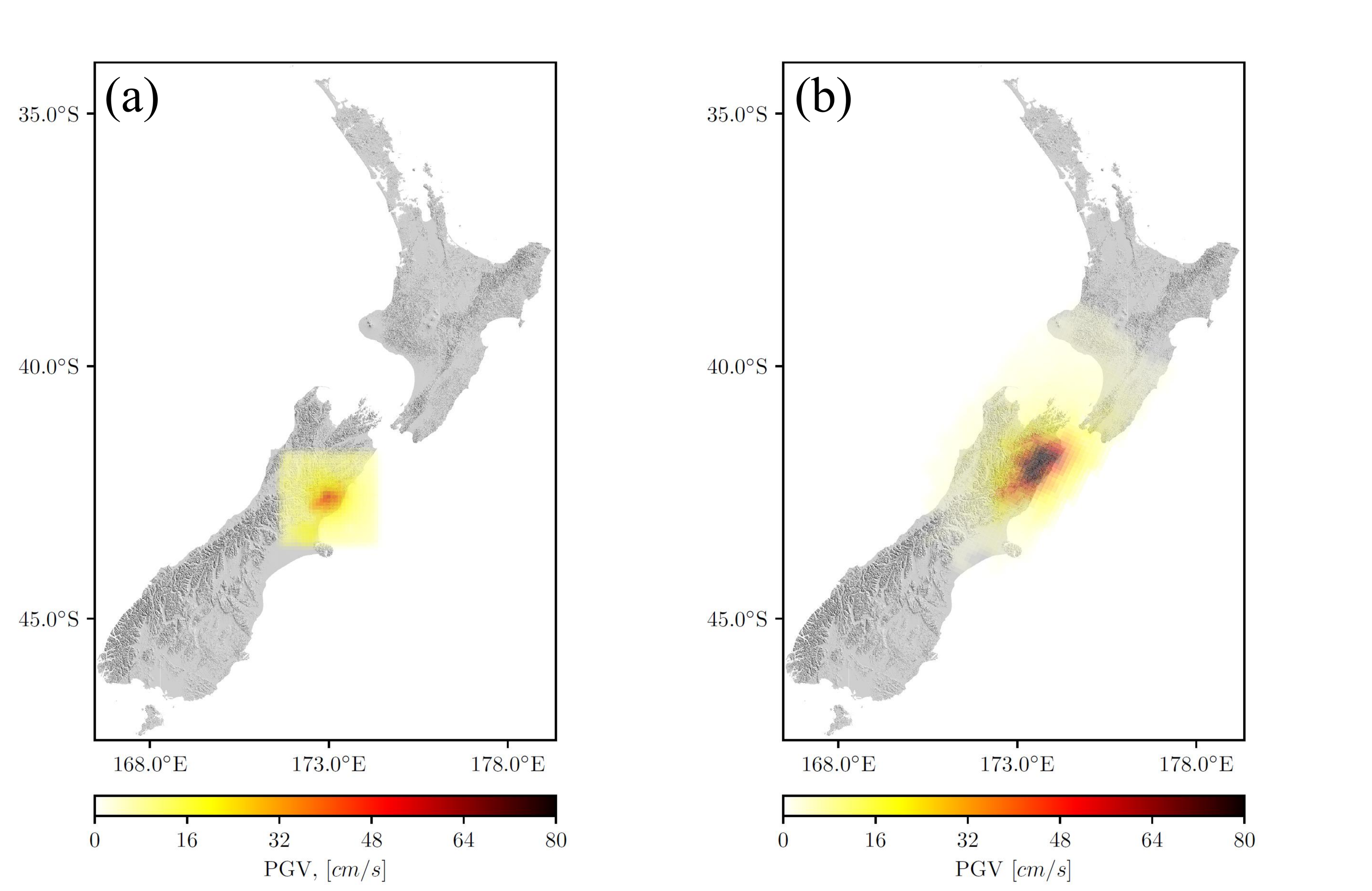

Figure 1 (on the right): Maps showing the Kaikoura M_w 7.8 earthquake GM intensity using (a) the point-source assumption obtained after a few hours following the event, and (b) the best estimate of the same event from Bradley et al. (2017) after a few months

Training data: New Zealand Cybershake v18.6

Cybershake is an ambitious program that aims to develop probabilistic hazard maps based on physics-based GM simulations. For each known New Zealand fault, a defined number of simulations are realized. Basic parameters of the simulations are selected via a Monte Carlo scheme, making each simulation unique. The ground motion intensity is recorded at existing instrument locations and over a country-wide uniform grid. The location of the hypocenter being of prime importance to determine the GM spatial distribution, simulations from the same faults are grouped together based on some carefully selected characteristics via K-mean clustering. As shown in Figure 2, a data point contains therefore: (1) the earthquake cluster ID, (2) the spatial distribution of the ground motion, and (3) the ground motion intensity at instrument locations. In total, about 17’000 ground motions have been simulated for 482 faults.

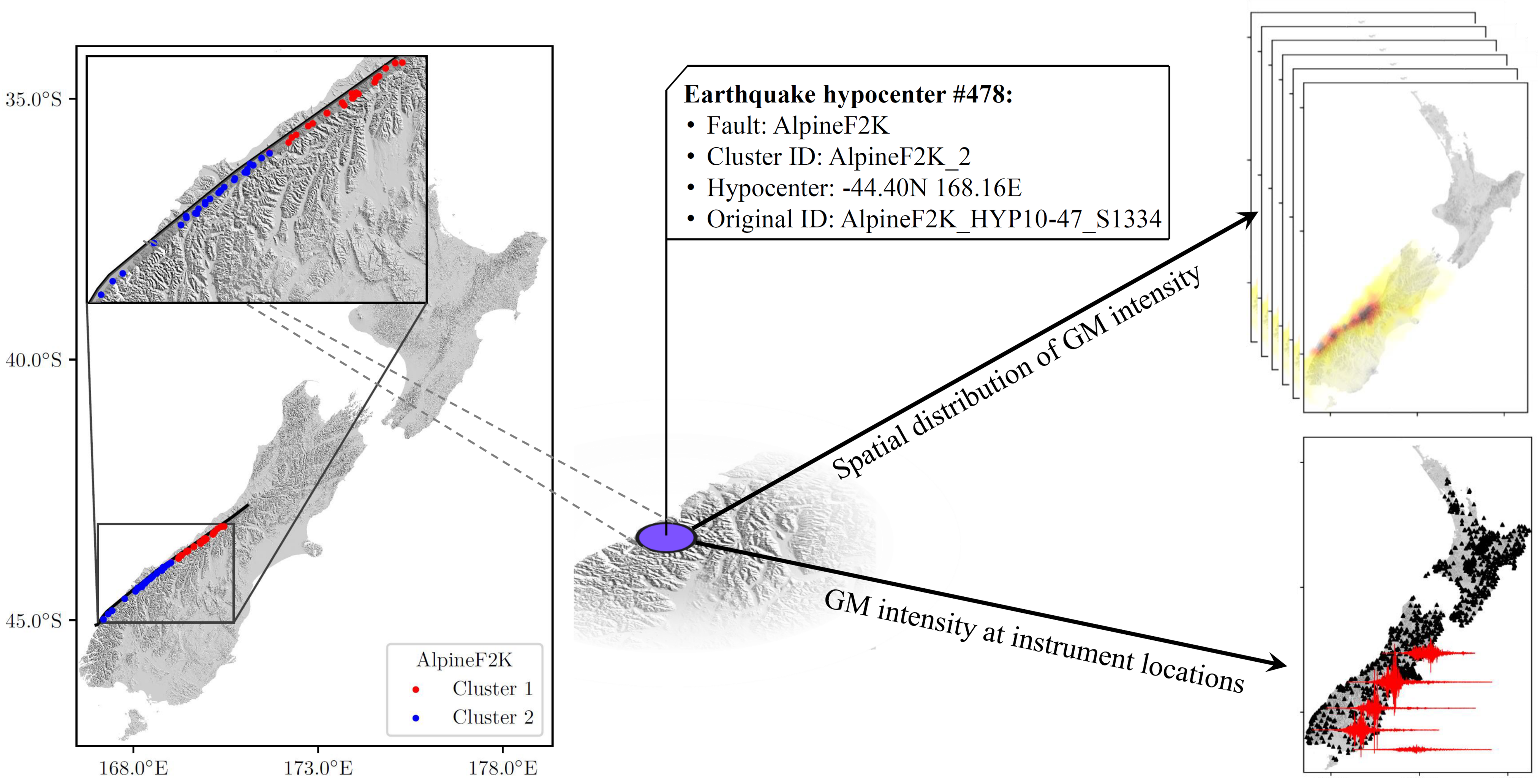

Figure 2: Example of a data point composed of an hypocenter from the AlpineF2K fault. (a) Map of the clustered hypocenters, (b) ground motion characteristics at the seismographs, and (c) multi-channel simulated ground motion map.

Overview of the method

In order to estimate the geospatial intensity of the earthquake, GM intensity captured at instrument locations X_i are first given to a random-forest-based rupture discriminator trained to determine the most likely earthquake source. The output of the random forest provides the probability of rupture P_j for each considered fault j. This vector of probability can be considered as a compressed signal of the geospatial GM intensity. A GM generator based on a deconvolution network is used to estimate the spatially distributed GM intensity.

The training of such a system is realized in two distinct steps: (1) the rupture discriminator is trained utilizing the simulated GM intensity at instrument locations as input and the rupture label as target, and (2) the GM generator is trained against the encoded GM intensity at instrument locations P_i as input and the simulated spatial distribution of the GM intensity.

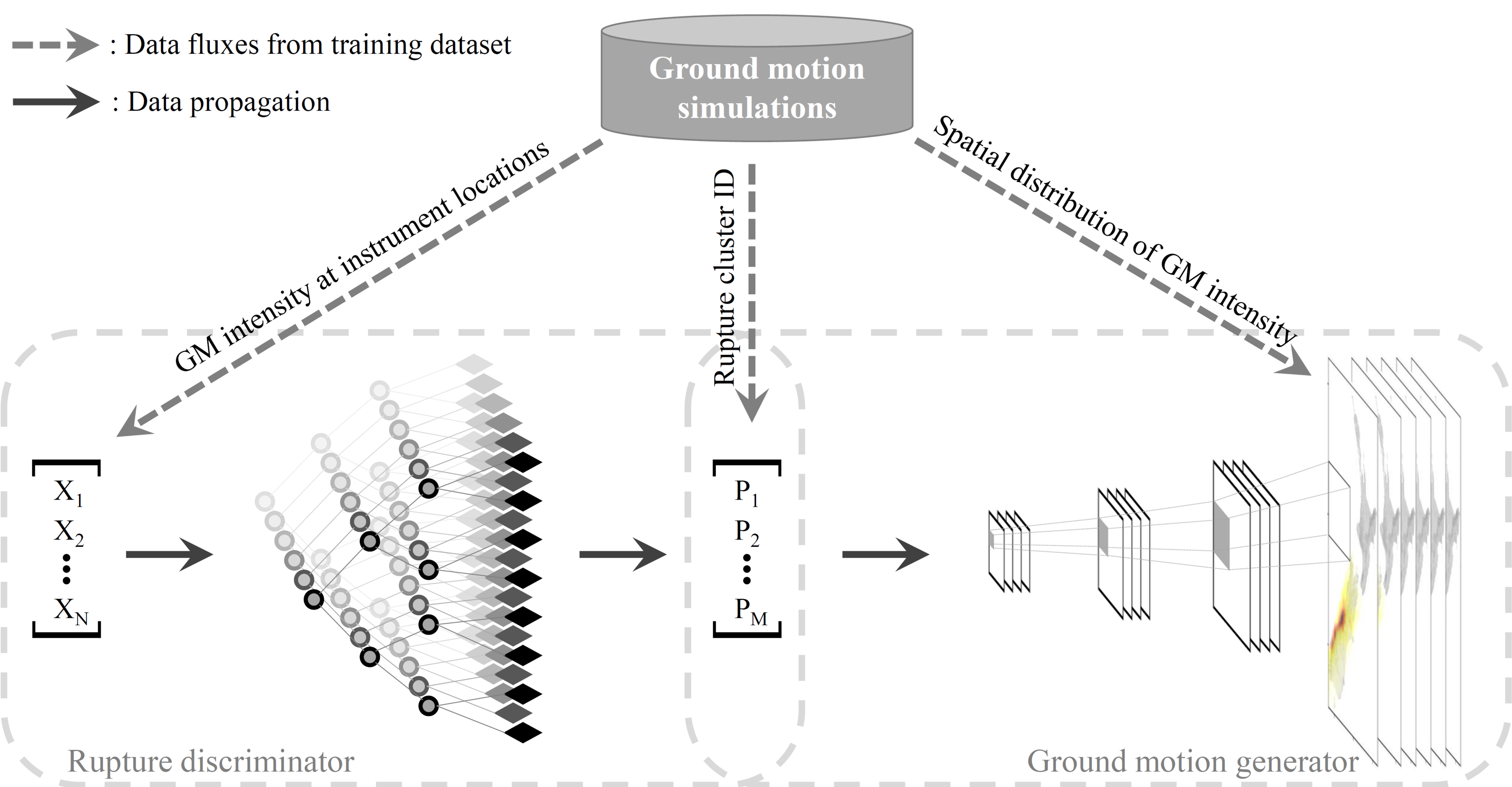

Figure 3: Overview of the method and its training. X_i: GM intensity at a particular instrument. P_j: probability that fault j has ruptured. GM intensity maps: estimated GM spatial distribution of different intensity measures.

GM generation results and future work

The trained GM generator is finally tested against some test data. Figure 4 shows an example of simulated (Fig. 4a) vs generated (Fig. 4b) GM intensity maps. Despite the overestimation that can be observed at the northern end of the rupture, results seem to be relatively well predicted by the generator. Most of the zones experiencing a ground motion susceptible to cause damage (PGV>20cm/s) remain within reasonable residual range (Fig. 4c).

Future work will focus on the ability to generate ground motion intensity maps for complex ruptures combining multiple sources. To achieve this, the training dataset will be composed of simulations where earthquake sources are selected in a combinatorial fashion and include data from smaller observed events. Earthquake sources will not be considered as finite anymore, but as a combination of cells.

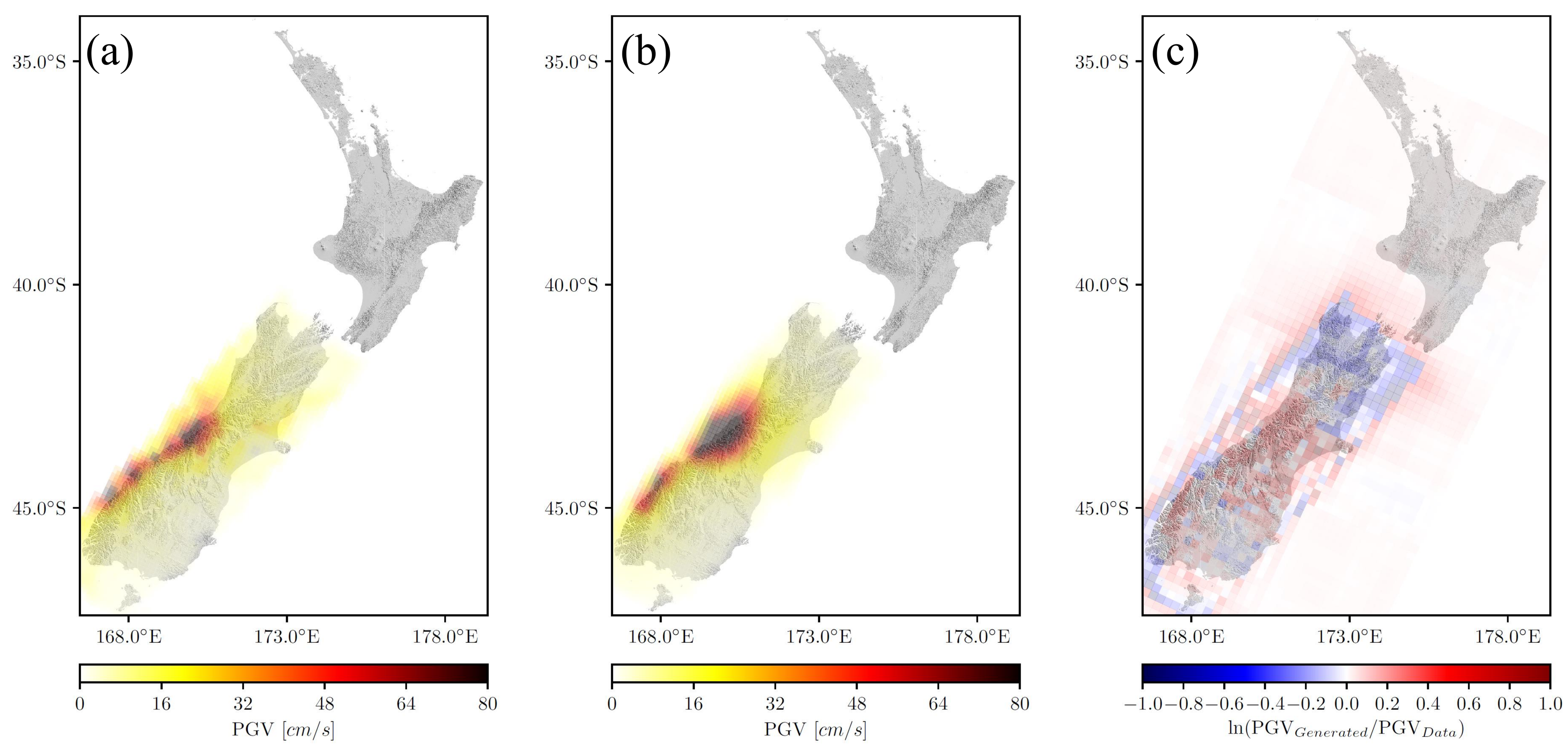

Figure 4 Example of (a) an AlpineF2K simulation, (b) its generated counterpart and (c) the residual between them